ABSTRACT OF THE DISCLOSURE

Disclosed are methods and apparatus for efficiently setting up and maintaining a defect classification system. In general terms, the setup procedure optionally includes automatically grouping a set of provided defects (*e.g.,* defect images) and presenting a representative set from each defect group to the user for classification. Alternatively, a representative set from the whole defect set may be presented to the user for classification without first grouping the defects into groups. The representative set does not include all of the defects and is selected to optimize manual classification efficiency. After the initial manual classification of the representative defects, the setup procedure includes an automatic procedure for classifying the non-reviewed or unclassified defects based on the manual class codes from the user-reviewed defects. After the automatic classification operation, the user may also be presented with defects from each class which may require re-classification. In particular embodiments, the user is iteratively presented with defects which have classifications that are suspect, which are near classification boundaries, or have classifications that have a low confidence level until each class is pure or contains a same type of defect classes as assigned by the user. After all the defects are manually classified through the above procedure, the classifier training set may be created automatically. New type of defects may be detected by using the existing training defects or the classified defects as the reference defects. The classifier maintenance may be done by merging existing classifiers together, adding new type of defects into the classifier, adding new boundary defects into the classifier, and removing redundant defects from the classifier.